

FIG. 1

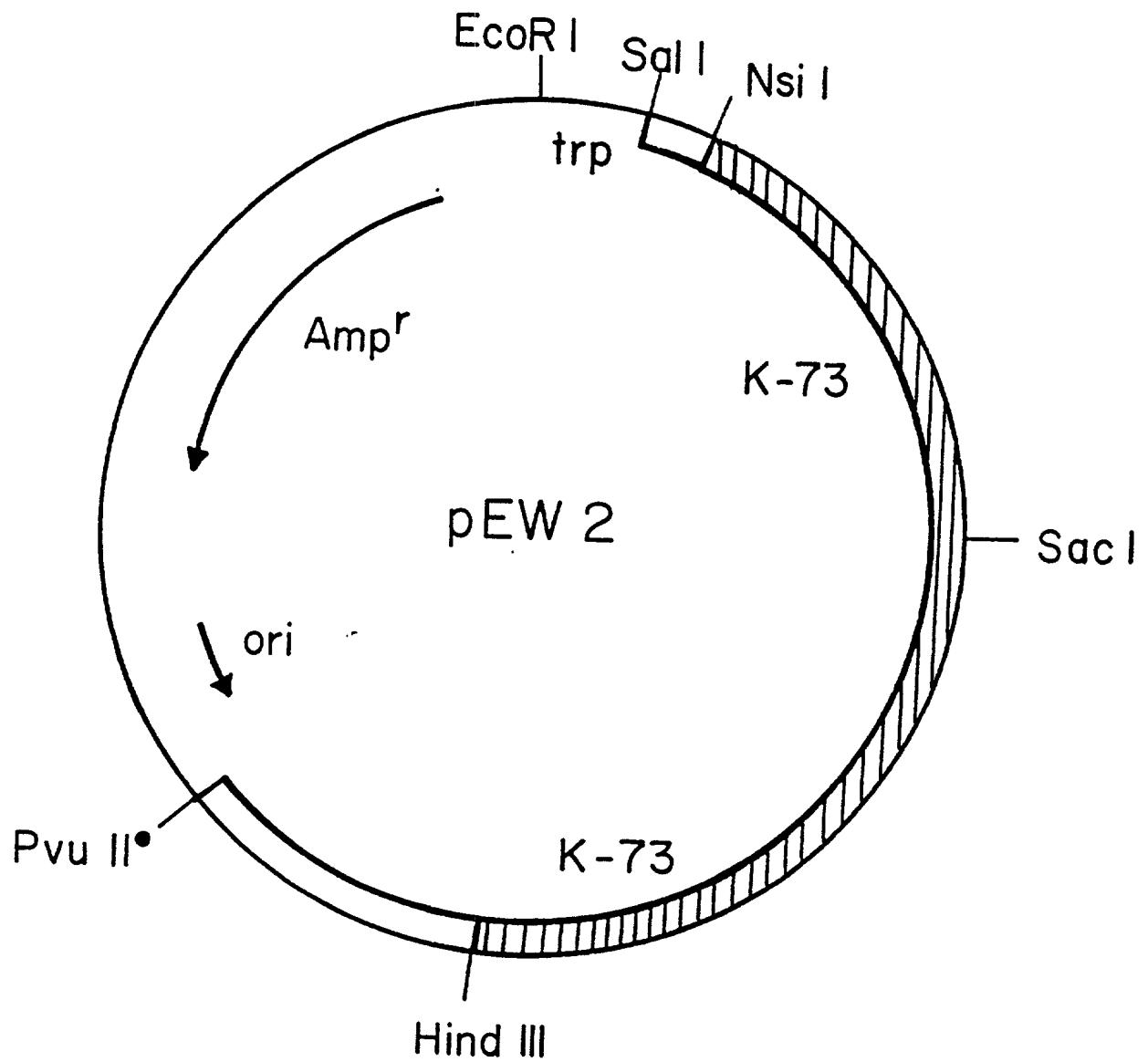


FIG. 2

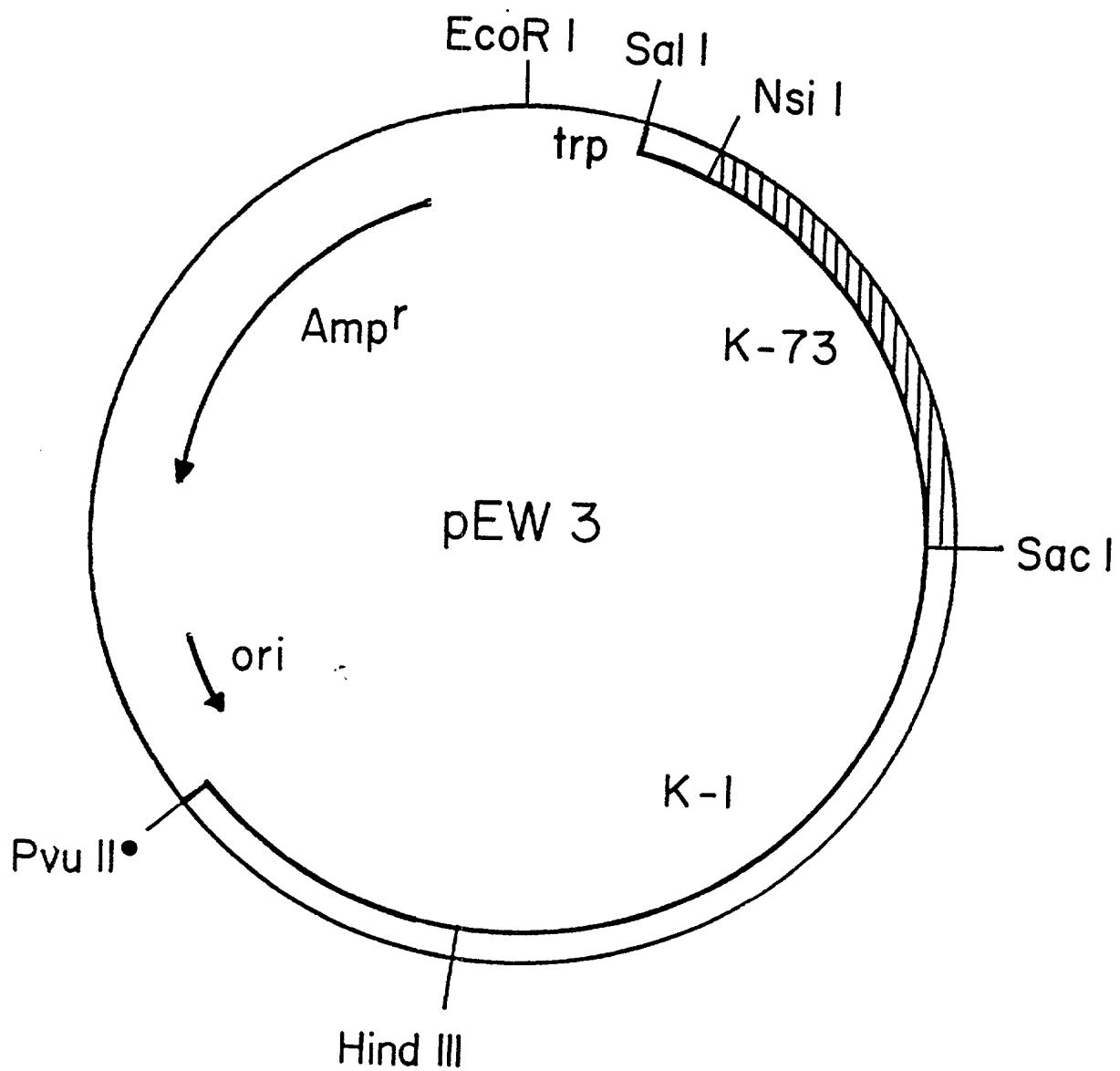


FIG. 3

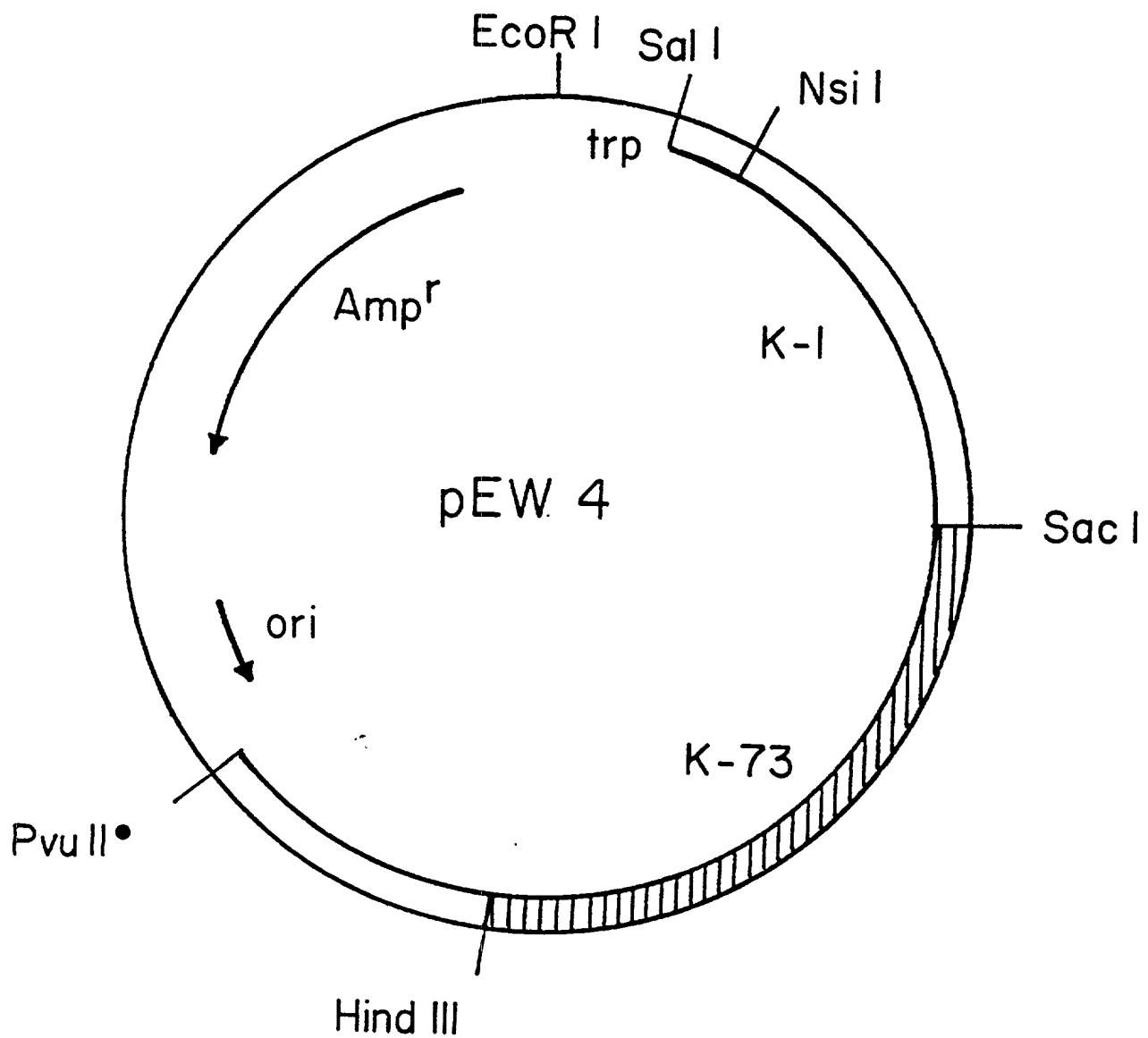


FIG. 4

(start HD-73) ATG GATAACAATC 400
 CGAACATCAA TGAATGCATT CCTTATAATT GTTTAAGTAA CCCTGAAGTA
 GAAGTATTAG GTGGAGAAAAG AATAGAAACT GGTACACCC CAATCGATAT 500
 TTCCCTTGTCG CTAACGCAAT TTCTTTGAG TGAATTTGTT CCCGGTGCTG
 GATTTGTGTT AGGACTAGTT GATATAATAT GGGGAATTT TGTCCTCT 600
 CAATGGGACG CATTCTTGT ACAAAATTGAA CAGTTAATTA ACCAAAGAAT
 AGAAGAATTG GCTAGGAACC AAGCCATTT TAGATTAGAA GGACTAAGCA 700
 ATCTTATCA AATTACGCA GAATCTTTA GAGAGTGGGA AGCAGATCCT
 ACTAATCCAG CATTAAAGAGA AGAGATGCGT ATTCAATTCA ATGACATGAA 800
 CAGTGCCCTT ACAACCGCTA TTCCCTTTTG TGCAAGTCAA AATTATCAAG
 TTCCCTTTTG ATCAGTATAT GTTCAAGCTG CAAATTTACA TTTATCAGTT 900
 TTGAGAGATG TTTCAGTGTT TGGACAAAGG TGGGGATTTG ATGCCGCGAC
 TATCAATAGT CGTTATAATG ATTTAACTAG GCTTATTGGC AACTATACAG 1000
 ATTATGCTGT ACGCTGGTAC AATACGGGAT TAGAACGTGT ATGGGGACCG
 GATTCTAGAG ATTGGGTAAG GTATAATCAA TTTAGAAGAG AATTAACACT 1100
 AACTGTATTA GATATCGTT CTCTGTTCCC GAATTATGAT AGTAGAAGAT
 ATCCAATTG AACAGTTTCC CAATTAACAA GAGAAATTAA TACAAACCCA 1200
 GTATTAGAAA ATTTGATGG TAGTTTCGA GGCTCGGCTC AGGGCATAGA
 AAGAAGTATT AGGAGTCCAC ATTTGATGGA TATACTAAC AGTATAACCA 1300
 TCTATACGGA TGCTCATAGG GGTTATTATT ATTGGTCAGG GCATCAAATA
 ATGGCTTCTC CTGTAGGGTT TTCGGGGCCA GAATTCACTT TTCCGCTATA 1400
 TGGAACTATG GGAAATGCAG CTCCACAAACA ACGTATTGTT GCTCAACTAG
 GTCAGGGCGT GTATAGAACAA TTATCGTCCA CTTTATATAG AAGACCTTT 1500
 AATATAGGGA TAAATAATCA ACAACTATCT GTTCTTGACG GGACAGAATT
 TGCTTATGGA ACCTCCTCAA ATTTGCCATC CGCTGTATAC AGAAAAAGCG 1600
 GAACGGTAGA TTCGCTGGAT GAAATACCGC CACAGAATAA CAACGTGCCA
 CCTAGGCAAG GATTTAGTCA TCGATTAAGC CATGTTCAA TGTTCGTTC 1700
 AGGCTTACTG AATAGTAGTG TAAGTATAAT AAGAGCT (end hd-73)
 (start HD-1) CCAACGT TTTCTTGGCA GCATCGCAGT 1900
 GCTGAATTAA ATAATATAAT TCCTTCATCA CAAATTACAC AAATACCTTT
 AACAAAATCT ACTAATCTTG GCTCTGGAAC TTCTGTCGTT AAAGGACCG 2000
 GATTTACAGG AGGAGATATT CTTCGAAGAA CTTCACCTGG CCAGATTCA
 ACCTTAAGAG TAAATATTAC TGCACCATTA TCACAAAGAT ATCGGGTAAG 2100
 AATTTCGCTAC GCTTCTACTA CAAATTACA ATTCCATACA TCAATTGACG
 GAAGACCTAT TAATCAGGGT AATTTTCAG CAACTATGAG TAGTGGGAGT 2200
 AATTTACAGT CGGGAABCTT TAGGACTGTA GGTTTACTA CTCCGTTAA
 CTTTCAAAT GGATCAAGTG TATTTACGTT AAGTGCTCAT GTCTTCAATT 2300
 CAGGCAATGA AGTTTATATA GATCGAATTG AATTTGTTCC GGCAGAAGTA
 ACCTTTGAGG CAGAATATGA TTTAGAAAGA GCACAAAAGG CGGTGAATGA 2400
 GCTGTTACT TCTTCCAATC AAATCGGGTT AAAACACAGAT GTGACGGATT
 ATCATATTGA TCAAGTATCC AATTTAGTTG AGTGTTCATC AGATGAATT 2500
 TGCTGGATG AAAAACAAAGA ATTGTCCGAG AAAGTCAAAC ATGCGAAGCG
 ACTTAGTGTGAT GAGCGGAATT TACTTCAAGA TCCAAACTTC AGAGGGATCA 2600
 ATAGACAACT AGACCGTGGC TGGAGAGGAA GTACGGATAT TACCATCCAA
 GGAGGCAGTG ACGTATTCAA AGAGAATTAC GTTACGCTAT TGGGTACCTT 2700
 TGATGAGTGC TATCCAACGT ATTTATATCA AAAATAGAT GAGTCGAAAT

FIG. 5A

bioRxiv preprint doi: <https://doi.org/10.1101/2022.07.05.500000>; this version posted July 6, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

TAAAAGCCTA TACCCGTTAT CAATTAAGAG GGTATATCGA AGATAGTCAA 2800
GACTTAGAAA TCTATTAAAT TCGCTACAAT GCAAAACATG AAACAGTAAA
TGTGCCAGGT ACGGGTTCCT TATGGCCGCT TTCAGCCCAA AGTCCAATCG 2900
GAAAGTGTGG AGAGCCGAAT CGATGCGCGC CACACCTTGA ATGGAATCCT
GACTTAGATT GTTCGTGTAG GGATGGAGAA AAGTGTGCC ATCATTGC 3000
TCATTTCTCC TTAGACATTG ATGTAGGATG TACAGACTTA AATGAGGACC
TAGGTGTATG GGTGATCTT AAGATTAAGA CGCAAGATGG GCACGCAAGA 3100
CTAGGGAATC TAGAGTTCT CGAAGAGAAA CCATTAGTAG GAGAAGCGCT
AGCTCGTGTG AAAAGAGCGG AGAAAAAAATG GAGAGACAAA CGTAAAAAT 3200
TGGAAATGGGA AACAAATATC GTTTATAAAG AGGCAAAAGA ATCTGTAGAT
GCTTATTG TAAACTCTCA ATATGATCAA TTACAAGCGG ATACGAATAT 3300
TGCCATGATT CATGCGGCAG ATAAACGTGT TCATAGCATT CGAGAAGCTT
ATCTGCCTGA GCTGTCTGTG ATTCCGGGTG TCAATGCGGC TATTTTGAA 3400
GAATTAGAAG GGCGTATTTT CACTGCATTG TCCCTATATG ATGCGAGAAA
TGTCTTTAA AATGGTGTGATT TTAATAATGG CTTATCCTGC TGGAACGTGA 3500
AAGGGCATGT AGATGTAGAA GAACAAAACA ACCAACGTTC GGTCCCTGTT
CTTCCGGAAT GGGAAAGCAGA AGTGTACCAA GAAGTTCGTG TCTGTCCGGG 3600
TCGTGGCTAT ATCCTTCGTG TCACAGCGTA CAAGGAGGG A TATGGAGAAG
GTTGCGTAAC CATTGATGAG ATCGAGAAACA ATACAGACGA ACTGAAGTTT 3700
AGCAACTGCG TAGAAGAGGA AATCTATCCA AATAAACACGG TAACGTGTAA
TGATTATACT GTAAATCAAG AAGAATAACGG AGGTGCGTAC ACTTCTCGTA 3800
ATCGAGGATA TAACGAAGCT CCTTCCGTAC CAGCTGATT TGCGTCAGTC
TATGAAGAAA AATCGTATAC AGATGGACGA AGAGAGAATC CTTGTGAATT 3900
TAACAGAGGG TATAGGGATT ACACGCCACT ACCAGTTGGT TATGTGACAA
AAGAATTAGA ATACTTCCC A GAAACCGATA AGGTATGGAT TGAGATTGGA 4000
GAAACCGGAAG GAACATTAT CGTGGACAGC GTGGATTAC TCCTTATGGGA
GGAA (end HD-1)

FIG. 5B

M D N N P N I N E C I P Y N C L S N P E V E V L G G E R I E
T G Y T P I D I S L S L T Q F L L S E F V P G A G F V L G L
V D I I W G I F G P S Q W D A F L V Q I E Q L I N Q R I E E
F A R N Q A I S R L E G L S N L Y Q I Y A E S F R E W E A D
P T N P A L R E E M R I Q F N D M N S A L T T A I P L F A V
Q N Y Q V P L L S V Y V Q A A N L H L S V L R D V S V F G Q
R W G F D A A T I N S R Y N D L T R L I G N Y T D Y A V R W
Y N T G L E R V W G P D S R D W V R Y N Q F R R E L T L T V
L D I V A L F P N Y D S R R Y P I R T V S Q L T R E I Y T N
P V L E N F D G S F R G S A Q G I E R S I R S P H L M D I L
N S I T I Y T D A H R G Y Y W S G H Q I M A S P V G F S G
P E F T F P L Y G T M G N A A P Q Q R I V A Q L G Q G V Y R
T L S S T L Y R R P F N I G I N N Q Q L S V L D G T E F A Y
G T S S N L P S A V Y R K S G T V D S L D E I P P Q N N N V
P P R Q G F S H R L S H V S M F R S G F S N S S V S I I R A
P T F S W Q H R S A E F N N I I P S S Q I T Q I P L T K S T
N L G S G T S V V K G P G F T G G D I L R R T S P G Q I S T
L R V N I T A P L S Q R Y R V R I R Y A S T T N L Q F H T S
I D G R P I N Q G N F S A T M S S G S N L Q S G S F R T V G
F T T P F N F S N G S S V F T L S A H V F N S G N E V Y I D
R I E F V P A E V T F E A E Y D L E R A Q K A V N E L F T S
S N Q I G L K T D V T D Y H I D Q V S N L V E C L S D E F C
L D E K Q E L S E K V K H A K R L S D E R N L L Q D P N F R
G I N R Q L D R G W R G S T D I T I Q G G D D V F K E N Y V
T L L G T F D E C Y P T Y L Y Q K I D E S K L K A Y T R Y Q
L R G Y I E D S Q D L E I Y L I R Y N A K H E T V N V P G T
G S L W P L S A Q S P I G K C G E P N R C A P H L E W N P D
L D C S C R D G E K C A H H S H H F S L D I D V G C T D L N
E D L G V W V I F K I K T Q D G H A R L G N L E F L E E K P
L V G E A L A R V K R A E K K W R D K R E K L E W E T N I V
Y K E A K E S V D A L F V N S Q Y D Q L Q A D T N I A M I H
A A D K R V H S I R E A Y L P E L S V I P G V N A A I F E E
L E G R I F T A F S L Y D A R N V I K N G D F N N G L S C W
N V K G H V D V E E Q N N Q R S V L V L P E W E A E V S Q E
V R V C P G R G Y I L R V T A Y K E G Y G E G C V T I H E I
E N N T D E L K F S N C V E E E I Y P N N T V T C N D Y T V
N Q E E Y G G A Y T S R N R G Y N E A P S V P A D Y A S V Y
E E K S Y T D G R R E N P C E F N R G Y R D Y T P L P V G Y
V T K E L E Y F P E T D K V W I E I G E T E G T F I V D S V
E L L L M E E

FIG. 6

(start HD-1) ATGG ATAACAATCC GAACATCAAT
 GAATGCATTC CTTATAATTG TTTAAGTAAC CCTGAAGTAG AAGTATTAGG 600
 TGGAGAAAGA ATAGAAAATG GTTACACCCCC AATCGATATT TCCTTGTCGC
 TAACGCAATT TCTTTTGAGT GAATTGTTT CCGGTGCTGG ATTTGTGTTA 700
 GGACTAGTTG ATATAATATG GGGAAATTTTT GGTCCCTCTC AATGGGACGC
 ATTTCTGTGAA CAAATTGAAAC AGTTAATTAA CCAAAAGAATA GAAGAATTG 800
 CTAGGAACCA AGCCATTCT AGATTAGAAG GACTAAGCAA TCTTATCAA
 ATTTACGCGAG AATCTTTAG AGAGTGGGAA GCAGATCCTA CTAATCCAGC 900
 ATTAAGAGAA GAGATGCGTA TTCAATTCAA TGACATGAAC AGTGCCTTA
 CAACCGCTAT TCCTCTTTG GCAGTTCAAA ATTATCAAGT TCCTCTTTA 1000
 TCAGTATATG TTCAAGCTGC AAATTTACAT TTATCAGTT TGAGAGATGT
 TTCAGTGTGTT GGACAAAGGT GGGGATTGGA TGCCGCGACT ATCAATAGTC 1100
 GTTATAATGA TTTAACCTAGG CTTATTGGCA ACTATACAGA TTATGCTGTG
 CGCTGGTACA ATACGGGATT AGAGCGTGTA TGGGGACCGG ATTCTAGAGA 1200
 TTGGGTAAGG TATAATCAAT TTAGAAGAGA GCTAACACTT ACTGTATTAG
 ATATCGTTGC TCTATTCTCA AATTATGATA GTCGAAGSTA TCCAATTGCA 1300
 ACAGTTTCCC AATTAACAAG AGAAAATTTAT ACGAACCCAG TATTAGAAAA
 TTTTGATGGT AGTTTTCGTG GAATGGCTCA GAGAATAGAA CAGAATATTA 1400
 GGCAACCACA TCTTATGGAT ATCCTTAATA GTATAACCAT TTATACTGAT
 GTGCATAGAG GCTTTAATTA TTGGTCAGGG CATCAAATAA CAGCTTCTCC 1500
 TGTAGGGTTT TCAGGACAG AATTGCGATT CCCTTTATTG GGGAAATGCGG
 GGAATGCGC TCCACCCGTA CTTGTCAT TAACGGTTT GGGGATTTT 1600
 AGAACATTAT CTTCACCTT ATATAGAAGA ATTATACCTG GTTCAGGCC
 AAATAATCAG GAACTGTTG TCCTTGATGG AACGGAGTTT TCTTTGCT 1700
 CCCTAACGAC CAACCTGCCT TCCACTATAT ATAGACAAAG GGGTACAGTC
 GATTCACTAG ATGTAATACC GCCACAGGAT AATAGTGTAC CACCTCGTGC 1800
 GGGATTTAGC CATCGATTGA GTCATGTTAC AATGCTGAGC CAAGCAGCTG
 GAGCAGTTA CACCTTGAGA GCTCAACGT (stop HD-1)
 (start HD-73) CCT ATGTTCTCTT
 GGATACATCG TAGTGCTGAA TTTAATAATA TAATTGCATC GGATAGTATT 1800
 ACTCAAATCC CTGCACTGAA GGGAAACTTT CTTTTAATG GTTCTGTAAT
 TTCAGGACCA GGATTTACTG GTGGGGACTT AGTTAGATTA AATAGTAGTG 1900
 GAAATAACAT TCAGAAATAGA GGGTATATTG AAGTTCCAAT TCACTTCCCA
 TCGACATCTA CCAGATATCG AGTTCGTGTA CGGTATGCTT CTGTAACCCC 2000
 GATTCACTC AACGTTAATT GGGGTAATTG ATCCATTGTT TCCAATACAG
 TACCAAGCTAC AGCTACGTCA TTAGATAATC TACAATCAAG TGATTTGGT 2100
 TATTTGAAA GTGCCAATGC TTTTACATCT TCATTAGGTA ATATAGTAGG
 TGTTAGAAAT TTTAGTGGGA CTGCAGGAGT GATAATAGAC AGATTTGAAT 2200
 TTATTCCAGT TACTGCAACA CTCGAGGCTG AATATAATCT GGAAAGAGCG

FIG. 7A

CAGAAGGC GG TGAATGCGCT GTTACGTCT ACAAAACCAAC TAGGGCTAAA 2300
AACAAATGTA ACGGATTATC ATATTGATCA AGTGTCCAAT TTAGTTACGT
ATTTATCGGA TGAATTGTCT CTGGATGAAA AGCGAGAATT GTCCGAGAAA 2400
GTCAAACATG CGAACGCGACT CAGTGATGAA CGCAATTAC TCCAAGATTC
AAATTTCAAA GACATTAATA GGCAACCAGA ACGTGGGTGG GGCGGAAGTA 2500
CAGGGATTAC CATCCAAGGA GGGGATGACG TATTTAAAGA AAATTACGTC
ACACTATCAG GTACCTTGA TGAGTGCTAT CCAACATATT TGTATCAAAA 2600
AATCGATGAA TCAAAATTAA AAGCCTTAC CCGTTATCAA TTAAGAGGGT
ATATCGAAGA TAGTCAAGAC TTAGAAATCT ATTTAATTG CTACAATGCA 2700
AAACATGAAA CAGTAAATGT GCCAGGTACG GGTTCTTAT GGCGCCTTTC
AGCCCCAAAGT CCAATCGGAA AGTGTGGAGA GCCGAATCGA TGCSCGCCAC 2800
ACCTTGAATG GAATCCTGAC TTAGATTGTT CGTGTAGGGA TGGAGAAAAAG
TGTGCCCATC ATTGCGATCA TTTCTCCTTA GACATTGATG TAGGATGTAC 2900
AGACTTAAAT GAGGACCTAG GTGTATGGGT GATCTTAAG ATTAAGACGC
AAGATGGGCA CGCAAGACTA GGGAAATCTAG AGTTTCTCGA AGAGAAACCA 3000
TTAGTAGGAG AAGCGCTAGC TCGTGTGAAA AGAGCGGAGA AAAAATGGAG
AGACAAACGT GAAAAATTGG AATGGGAAAC AAATATCGTT TATAAAGAGG 3100
CAAAAGAATC TGTAGATGCT TTATTTGTAA ACTCTCAATA TGATCAATT
CAAGCGGATA CGAATATTGC CATGATTCTAT GCGGCAGATA AACGTGTTCA 3200
TAGCATTGCA GAAGCTTATC TGCTGAGCT GTCTGTGATT CGGGGTGTCA
ATGCGCTAT TTTTGAAGAA TTAGAAGGGC GTATTTCAC TGCATTCTCC 3300
CTATATGATG CGAGAAATGT CATTAAAAAT GGTGATTTA ATAATGGCTT
ATCCTGCTGG AACGTGAAAG GGCATGTAGA TGTAGAAGAA CAAAACAACC 3400
AACGTTCGGT CCTTGTTGTT CGGAATGGG AAGCAGAAGT GTCACAAGAA
GTTCGTGTCT GTCCGGGTG 76 TGGCTATATC CTTCGTGTCA CAGCGTACAA 3500
GGAGGGATAT GGAGAAGGTT GCGTAACCCT TCATGAGATC GAGAACAAATA
CAGACGAACT GAAGTTTAGC AACTGCGTAG AAGAGGAAAT CTATCCAAT 3600
AACACGGTAA CGTGTAAATGA TTATACTGTA AATCAAGAAG AATACGGAGG
TGCCTACACT TCTCGTAATC GAGGATATAA CGAAGCTCCT TCCGTACCAAG 3700
CTGATTATGC GTCAGTCTAT GAAGAAAAAT CGTATACAGA TGGACGAAAGA
GAGAACCTT GTGAATTAA CAGAGGGTAT AGGGATTACA CGCCACTACC 3800
AGTTGGTTAT GTGACAAAAAG AATTAGAATA CTTCCCAGAA ACCGATAAGG
TATGGATTGA GATTGGAGGA ACGGAAGGAA CATTATCGT GGACAGCGTG 3900
GAATTACTCC TTATGGAGGA A (end HD-73)

FIG. 7B

M D N N P N I N E C I P Y N C L S N P E V E V L G G E R I E
T G Y T P I D I S L S L T Q F L L S E F V P G A G F V L G L
V D I I W G I F G P S Q W D A F P V Q I E Q L I N Q R I E E
F A R N Q A I S R L E G L S N L Y Q I Y A E S F R E W E A D
P T N P A L R E E M R I Q F N D M N S A L T T A I P L L A V
Q N Y Q V P L L S V Y V Q A A N L H L S V L R D V S V F G Q
R W G F D A A T I N S R Y N D L T R L I G N Y T D Y A V R W
Y N T G L E R V W G P D S R D W V R Y N Q F R R E L T L T V
L D I V A L F S N Y D S R R Y P I R T V S Q L T R E I Y T N
P V L E N F D G S F R G M A Q R I E Q N I R Q P H L M D I L
N S I T I Y T D V H R G F N Y W S G H Q I T A S P V G F S G
P E F A F P L F G N A G N A A P P V L V S L T G L G I F R T
L S S P L Y R R I I L G S G P N N Q E L F V L D G T E F S F
A S L T T N L P S T I Y R Q R G T V D S L D V I P P Q D N S
V P P R A G F S H R L S H V T M L S Q A A G A V Y T L R A Q
R P M F S W I H R S A E F N N I I A S D S I T Q I P A V K G
N F L F N G S V I S G P G F T G G D L V R L N S S G N N I Q
N R G Y I E V P I H F P S T S T R Y R V R V R Y A S V T P I
H L N V N W G N S S I F S N T V P A T A T S L D N L Q S S D
F G Y F E S A N A F T S S L G N I V G V R N F S G T A G V I
I D R F E F I P V T A T L E A E Y N L E R A Q K A V N A L F
T S T N Q L G L K T N V T D Y H I D Q V S N L V T Y L S D E
F C L D E K R E L S E K V K H A K R L S D E R N L L Q D S N
F K D I N R Q P E R G W G G S T G I T I Q G G D D V F K E N
Y V T L S G T F D E C Y P T Y L Y Q K I D E S K L K A F T R
Y Q L R G Y I E D S Q D L E I Y L I R Y N A K H E T V N V P
G T G S L W P L S A Q S P I G K C G E P N R C A P H L E W N
P D L D C S C R D G E K C A H H S H H F S L D I D V G C T D
L N E D L G V W V I F K I K T Q D G H A R L G N L E F L E E
K P L V G E A L A R V K R A E K K W R D K R E K L E W E T N
I V Y K E A K E S V D A L F V N S Q Y D Q L Q A D T N I A M
I H A A D K R V H S I R E A Y L P E L S V I P G V N A A I F
E E L E G R I F T A F S L Y D A R N V I K N G D F N N G L S
C W N V K G H V D V E E Q N N Q R S V L V V P E W E A E V S
Q E V R V C P G R G Y I L R V T A Y K E G Y G E G C V T I H
E I E N N T D E L K F S N C V E E E I Y P N N T V T C N D Y
T V N Q E E Y G G A Y T S R N R G Y N E A P S V P A D Y A S
V Y E E K S Y T D G R R E N P C E F N R G Y R D Y T P L P V
G Y V T K E L E Y F P E T D K V W I E I G E T E G T F I V D
S V E L L L M E E

FIG. 8

(start HD-73)

CGAACATCAA	TGAATGCATT	CCTTATAATT	GTTTAAGTAA	CCCTGAAGTA	ATG GATAACAATC	400
GAAGTATTAG	GTGGAGAAAAG	AATAGAAAACT	GGTTACACCC	CAATCGATAT		500
TTCCTTGTCG	CTAACGCAAT	TTCTTTGAG	TGAATTGTT	CCCGGTGCTG		
GATTTGTGTT	AGGACTAGTT	GATATAATAT	GGGGAATTTT	TGGTCCCTCT		600
CAATGGGACG	CATTTCTTGT	ACAAATTGAA	CAGTTAATTA	ACCAAAGAAT		
AGAAGAATTC	GCTAGGAACC	AAGCCATTTC	TAGATTAGAA	GGACTAAGCA		700
ATCTTATCA	AATTACGCA	GAATCTTTA	GAGAGTGGGA	AGCAGATCCT		
ACTAATCCAG	CATTAAGAGA	AGAGATGCGT	ATTCAATTCA	ATGACATGAA		800
CAGTGCCCTT	ACAACCGCTA	TTCCCTTTT	TGCAGTTCAA	AATTATCAAG		
TTCCCTTTT	ATCAGTATAT	GTTCAAGCTG	CAAATTACA	TTTATCAGTT		900
TTGAGAGATG	TTTCAGTGTG	TGGACAAAGG	TGGGGATTG	ATGCCGCGAC		
TATCAATAGT	CGTTATAATG	ATTTAACTAG	GCTTATTGGC	AACTATACAG		1000
ATTATGCTGT	ACGCTGGTAC	AATACGGGAT	TAGAACGTGT	ATGGGGACCG		
GATTCTAGAG	ATGGGTAAG	GTATAATCAA	TTTAGAAGAG	AATTAAACACT		1100
AACTGTATTA	GATATCGTTG	CTCTGTTCCC	GAATTATGAT	AGTAGAAGAT		
ATCCAATTG	AACAGTTTCC	CAATTAAACAA	GAGAAATTAA	TACAAACCCA		1200
GTATTAGAAA	ATTTTGATGG	TAGTTTCGA	GGCTCGGCTC	AGGGCATAGA		
AAGAAGTATT	AGGAGTCCAC	ATTTGATGGA	TATACTTAAC	AGTATAACCA		1300
TCTATACGGA	TGCTCATAGG	GGTTATTATT	ATTGGTCAGG	GCATCAAATA		
ATGGCTTCTC	CTGTAGGGTT	TTCGGGGCCA	GAATTCACTT	TTCCGCTATA		1400
TGGAACATATG	GGAAATGCAG	CTCCACAAACA	ACGTATTGTT	GCTCAACTAG		
GTCAGGGCGT	GTATAGAACAA	TTATCGTCCA	CTTTATATAG	AAGACCTTTT		1500
AATATAGGGA	TAAATAATCA	ACAACATATCT	GTTCCTGACG	GGACAGAATT		
TGCTTATGGA	ACCTCCTCAA	ATTTGCCATC	CGCTGTATAC	AGAAAAAGCG		1600
GAACGGTAGA	TTCGCTGGAT	GAAATACCGC	CACAGAATAA	CAACGTGCCA		
CCTAGGCAAG	GATTTAGTCA	TCGATTAAGC	CATGTTCAA	TGTTTCGTT		1700
AGGCTTTAGT	AATAGTAGTG	TAAGTATAAT	AAGAGCT	(end hd-73)		
(start HD-1)						
GCTGAATTAA	ATAATATAAT	TCCTTCATCA	CAAATTACAC	AAATACCTT	CCAACACGT	1900
AACAAAATCT	ACTAACTTGT	GCTCTGGAAC	TTCTGTCGTT	AAAGGACCG	TTCTGGCA	2000
GATTTACAGG	AGGAGATATT	CTTCGAAGAA	CTTCACCTGG	CCAGATTTC	GCATCGCAGT	
ACCTTAAGAG	TAATATTAC	TGCACCATTA	TCACAAAGAT	ATCGGGTAAG		2100
AATTGCTAC	GCTTCTACTA	CAAATTACA	ATTCCATACA	TCAATTGACG		
GAAGACCTAT	TAATCAGGGT	AATTTTCAG	CAACTATGAG	TAGTGGGAGT		2200
AATTTACAGT	CCGGAAGCTT	TAGGACTGTA	GGTTTTACTA	CTCCGTTAA		
CTTTCAAT	GGATCAAGTG	TATTTACGTT	AAAGTGCAT	GTCTTCATT		2300
CAGGCAATGA	AGTTTATATA	GATCGAATTG	AATTTGTTCC	GGCAGAAGTA		
ACCTTTGAGG	CAGAATATGA	TTTAGAAAAGA	GCACAAAAGG	CGGTGAATGA		2400
GCTGTTTACT	TCTTCCAATC	AAATCGGGTT	AAAAACAGAT	GTGACGGATT		
ATCATATTGA	TCAAGTATCC	AATTTAGTTG	AGTGTGTTATC	AGATGAATT		2500
TGTCTGGATG	AAAAACAAAGA	ATTGTCCGAG	AAAGTCAAAC	ATGCGAAGCG		
ACTTAGTGTG	GAGCGGAATT	TACTTCAAGA	TCCAAACTTC	AGAGGGATCA		2600
ATAGACAAC	AGACCGTGCG	TGGAGAGGAA	GTACGGATAT	TACCATCCAA		

FIG. 9A

GGAGGGCGATG ACGTATTCAA AGAGAATTAC GTTACGCTAT TGGGTACCTT 2700
TGATGAGTGC TATCCAACGT ATTTATATCA AAAAATAGAT GAGTCGAAAT
TAAAAGCCTA TACCCGTTAT CAATTAAGAG GGTATATCGA AGATAGTCAA 2800
GACTTAGAAA TCTATTTAAT TCGCTACAAT GCAAAACATG AAACAGTAAA
TGTGCCAGGT ACGGGTTCCCT TATGGCCGCT TTCAGCCCCA AGTCCAATCG 2900
GAAAGTGTGG AGAGCCGAAT CGATGCGCGC CACACCTTGA ATGGAATCCT
GACTTAGATT GTTCGTGTAG GGATGGAGAA AAGTGTGCC 3000 ATCATTCGCA
TCATTTCTCC TTAGACATTG ATGTAGGATG TACAGACTTA AATGAGGACC
TAGGTGTATG GGTGATCTTT AAGATTAAGA CGCAAGATGG GCACGCAAGA 3100
CTAGGGAATC TAGAGTTTCT CGAAGAGAAA CCATTAGTAG GAGAAGCGCT
AGCTCGTGTG AAAAGAGCGG AGAAAAAAATG GAGAGACAAA CGTGAAAAAT 3200
TGGAATGGGA AACAAATATC GTTTATAAAG AGGCAAAAGA ATCTGTAGAT
GCTTTATTTG TAAACTCTCA ATATGATCAA TTACAAGCGG ATACGAATAT 3300
TGCCATGATT CATGCAGGCAG ATAAACGTGT TCATAGCATT CGAGAACGTT
ATCTGCCTGA GCTGTCTGTG ATTCCGGGTG TCAATGCAGGC TATTTTGAA 3400
GAATTAGAAG GGC GTATTTC CACTGCATTG TCCCTATATG ATGCAGAGAAA
TGTCA TAAA AATGGTGATT TTAATAATGG CTTATCCTGC TGGAACGTGA 3500
AAGGGCATGT AGATGTAGAA GAACAAAACA ACCAACGTTG GGTCCCTTGT
CTTCCGGAAT GGGAAAGCAGA AGTGTACCAA GAAGTTCGTG TCTGTCCGGG 3600
TCGTGGCTAT ATCCTTCGTG TCACAGCGTA CAAGGAGGGG TATGGAGAAG
GTTGCSTAAC CATTCA TGAG ATCGAGAACAA ATACAGACGA ACTGAAGTTT 3700
AGCAACTGCG TAGAAGAGGA AATCTATCCA AATAACACGG TAACGTGTAA
TGATTATACT GTAAATCAAG AAGAATACGG AGGTGCCTAC ACTTCTCGTA 3800
ATCGAGGATA TAACGAAGCT CCTTCCGTAC CAGCTGATTA TGCGTCAGTC
TATGAAGAAA AATCGTATAAC AGATGGACGA AGAGAGAACATC CTTGTGAATT 3900
TAACAGAGGG TATAGGGATT ACACGCCACT ACCAGTTGGT TATGTGACAA
AAGAATTAGA ATACTTCCCA GAAACCGATA AGGTATGGAT TGAGATTGGA 4000
GAAACGGAAAG GAACATTAT CGTGGACAGC GTGGAAATTAC TCCTTATGGA
GGAA (end HD-1)

FIG. 9B

M D N N P N I N E C I P Y N C L S N P E V E V L G G E R I E
 T G Y T P I D I S L S L T Q F L L S E F V P G A G F V L G L
 V D I I W G I F G P S Q W D A F L V Q I E Q L I N Q R I E E
 F A R N Q A I S R L E G L S N L Y Q I Y A E S F R E W E A D
 P T N P A L R E E M R I Q F N D M N S A L T T A I P L F A V
 Q N Y Q V P L L S V Y V Q A A N L H L S V L R D V S V F G Q
 R W G F D A A T I N S R Y N D L T R L I G N Y T D Y A V R W
 Y N T G L E R V W G P D S R D W V R Y N Q F R R E L T L T V
 L D I V A L F P N Y D S R R Y P I R T V S Q L T R E I Y T N
 P V L E N F D G S F R G S A Q G I E R S I R S P H L M D I L
 N S I T I Y T D A H R G Y Y Y W S G H Q I M A S P V G F S G
 P E F T F P L Y G T M G N A A P Q Q R I V A Q L G Q G V Y R
 T L S S T L Y R R P F N I G I N N Q Q L S V L D G T E F A Y
 G T S S N L P S A V Y R K S G T V D S L N E I P P Q N N N V
 P P R Q E F S H R L S H V S M F R S G F S N S S V S I I R A
 P T F S W Q H R S A E F N N I I P S S Q I T Q I P L T K S T
 N L G S G T S V V K G P G F T G G D I L R R T S P G Q I S T
 L R V N I T A P L S Q R Y R V R I R Y A S T T N L Q F H T S
 I D G R P I N Q G N F S A T M S S G S N L Q S G S F R T V G
 F T T P F N F S N G S S V F T L S A H V F N S G N E V Y I D
 R I E F V P A E V T F E A E Y D L E R A Q K A V N E L F T S
 S N Q I G L K T D V T D Y H I D Q V S N L V E C L S D E F C
 L D E K Q E L S E K V K H A K R L S D E R N L L Q D P N F R
 G I N R Q L D R G W R G S T D I T I Q G G D D V F K E N Y V
 T L L G T F D E C Y P T Y L Y Q K I D E S K L K A Y T R Y Q
 L R G Y I E D S Q D L E I Y L I R Y N A K H E T V N V P G T
 G S L W P L S A Q S P I G K C G E P N R C A P H L E W N P D
 L D C S C R D G E K C A H H S H H F S L D I D V G C T D L N
 E D L G V W V I F K I K T Q D G H A R L G N L E F L E E K P
 L V G E A L A R V K R A E K K W R D K R E K L E W E T N I V
 Y K E A K E S V D A L F V N S Q Y D Q L Q A D T N I A M I H
 A A D K R V H S I R E A Y L P E L S V I P G V N A A I F E E
 L E G R I F T A F S L Y D A R N V I K N G D F N N G L S C W
 N V K G H V D V E E Q N N Q R S V L V L P E W E A E V S Q E
 V R V C P G R G Y I L R V T Á Y K E G Y G E G C V T I H E I
 E N N T D E L K F S N C V E E E I Y P N N T V T C N D Y T V
 N Q E E Y G G A Y T S R N R G Y N E A P S V P A D Y A S V Y
 E E K S Y T D G R R E N P C E F N R G Y R D Y T P L P V G Y
 V T K E L E Y F P E T D K V W I E I G E T E G T F I V D S V
 E L L L M E E

FIG. 10

(start HD-73) ATG GATAACAATC 400
 CGAACATCAA TGAATGCATT CCTTATAATT GTTTAAGTAA CCCTGAAGTA
 GAAGTATTAG CTGGAGAAAG AATAGAAAATC GGTTACACCC CAATCGATAT 500
 TTCCCTTGTGCG CTAACGCAAT TTCTTTGAG TGAATTGTT CCCGGTGCTG
 GATTGTTGTT AGGACTAGTT GATATAATAT GGGGAATTTC TGTTCCCTCT 600
 CAATGGGACG CATTCTTGT ACAAAATTGAA CAGTTAATTA ACCAAAGAAT
 AGAAGAATTG GCTAGGAACC AAGCCATTTC TAGATTAGAA GGACTAAGCA 700
 ATCTTATCA AATTTACGCA GAATCTTTA GAGAGTGGGA AGCAGATCCT
 ACTAATCCAG CATTAAAGAGA AGAGATGCGT ATTCAATTCA ATGACATGAA 800
 CABTGCCCCCT ACAACCGCTA TTCTCTTT TGCAAGTCAA AATTATCAAG
 TTCCCTTTT ATCAAGTATAT GTTCAGTGTG CAAATTACA TTTATCAGTT 900
 TTGAGAGATG TTTCAAGTGTG TGGACAAAGG TGGGGATTG ATGCCGCGAC
 TATCAATAGT CGTTATAATG ATTTAACTAG GCTTATTGGC AACTATAACAG 1000
 ATTATGCTGT ACGCTGGTAC AATAACGGGAT TAGAACGTGT ATGGGGACCG
 GATTCTAGAG ATTGGGTAAG GTATAATCAA TTAGAACAG AATTAACACT 1100
 AACTGTTAATGATATCA GATATCGTTG CTCTGTTCCC GAATTATGAT AGTAGAAGAT
 ATCCAATTG AACAGTTCC CAATTAACAA GAGAAATTAA TACAAACCCA 1200
 GTATTAGAAA ATTGATGG TAGTTTCGA GGCTCGGCTC AGGGCATAGA
 AAGAAGTATT AGGAGTCCAC ATTTGATGGA TATACTTAAC AGTATAACCA 1300
 TCTATACGGA TGCTCATAGG GGTTATTATT ATTGGTCAGG GCATCAAATA
 ATGGCTTCTC CTGTAGGGTT TTGGGGGCCA GAATTCACTT TTCCGCTATA 1400
 TGGAACATATG GGAAATGCAG CTCCACAACA ACGTATTGTT GCTCAACTAG
 GTCAGGGCGT GTATAGAACAA TTATCGTCCA CTTTATATAG AAGACCTTT 1500
 AATATAGGGA TAAATAATCA ACAACTATCT GTTCTTGACG GGACAGAATT
 TGCTTATGGA ACCTCCTCAA ATTGCCATC CGCTGTATAC AGAAAAAGCG 1600
 GAACGGTAGA TTCGCTGGAT GAAATACCGC CACAGAATAA CAACGTGCCA
 CCTAGGCAAG GATTAGTCA TCGATTAAGC CATGTTCAA TGTTTCGTTC 1700
 AGGCTTGTG AATAGTAGTG TAAGTATAAT AAGAGCT (end hd-73)
 (start HD-1) CCAACGT TTTCTGGCA GCATCGCAGT 1900
 GCTGAATTAA ATAATATAAT TCCTTCATCA CAAATTACAC AAATACCTT
 AACAAATCT ACTAATCTG GCTCTGGAAC TTCTGTCGTT AAAGGACCAG 2000
 GATTACAGG AGGAGATATT CTTCGAAGAA CTTCACCTGG CCAGATTCA
 ACCTTAAGAG TAAATATTAC TGCAACATTA TCACAAAGAT ATCGGGTAAG 2100
 AATTGCTTAC GCTTCTACTA CAAATTACA ATTCCATACA TCAATTGACG
 GAAGACCTAT TAATCAGGGT AATTTCAG CAACTATGAG TAGTGGGAGT 2200
 AATTACAGT CGGAAAGCTT TAGGACTGTA GGTTTACTA CTCCGTTAA
 CTTTCAAAT GGATCAAGTG TATTACGTT AAGTGCTCAT GTCTTCAATT 2300
 CAGGCAATGA AGTTTATATA GATCGAATTG AATTGTTCC GGCAAGAGTA
 ACCTTGAGG CAGAATATGA TTTAGAAAGA GCACAAAAGG CGGTGAATGA 2400
 GCTGTTACT TCTTCAAATC AAATCGGGTT AAAACAGAT GTGACGGATT
 ATCATATTGA TCAAGTATCC AATTAGTTG AGTGTTCATC AGATGAATT 2500
 TGTCTGGATG AAAAACAAAGA ATTGTCCGAG AAAGTCAAAC ATGCGAAGCG
 ACTTAGTGAT GAGCGGAATT TACTTCAAGA TCCAAACTTC AGAGGGATCA 2600
 ATAGACAACG AGACCGTGCG TGGAGAGGAA GTACGGATAT TACCATCAA
 GGAGGCGATG ACGTATTCAA AGAGAATTAC GTTACGCTAT TGGGTACCTT 2700
 TGATGAGTGC TATCCAACGT ATTTATATCA AAAATAGAT GAGTCGAAAT

FIG. 11A

TAAAAGCCTA TACCCGTTAT CAATTAAGAG GGTATATCGA AGATAAGTCAA 2800
GACTTAGAAA TCTATTTAAC TCGCTACAAT GCAAAACATG AAACAGTAAA
TGTGCCAGGT ACGGGTTCCCT TATGGCCGCT TTCAGCCCAA AGTCCAATCG 2900
GAAAGTGTGG AGAGCCGAAT CGATGCAGCG CACACCTTGA ATGGAATCCT
GACTTAGATT GTTCGTGTAG GGATGGAGAA AAGTGTGCC ATCATTCGCA 3000
TCATTTCTCC TTAGACATTG ATGTAGGATG TACAGACTTA AATGAGGACC
TAGGTGTATG GGTGATCTT AAGATTAAGA CGCAAGATGG GCACGCAAGA 3100
CTAGGGAATC TAGAGTTTCT CGAAGAGAAA CCATTAGTAG GAGAAGCGCT
AGCTCGTGTG AAAAGAGCGG AGAAAAAAATG GAGAGACAAA CGTGAAAAAT 3200
TGGAATGGGA AACAAATATC GTTTATAAAG AGGCAAAAGA ATCTGTAGAT
GCTTTATTTG TAAACTCTCA ATATGATCAA TTACAAGCGG ATACGAATAT 3300
TGCCATGATT CATGCAGCG ATAAACGTGT TCATAGCATT CGAGAAGCTT
ATCTGCCTGA GCTGTCGTG ATTCCGGGTG TCAATGCGGC TATTTTGAA 3400
GAATTAGAAG GGCGTATTT CACTGCATTG TCCCTATATG ATGCGAGAAA
TGTCAATTAA AATGGTGATT TTAATAATGG CTTATCCTGC TGGAACGTGA 3500
AAGGGCATGT AGATGTAGAA GAACAAAACA ACCAACGTTG GGTCTTGTGTT
CTTCCCGGAAT GGGAAAGCAGA AGTGTACCAA GAAGTTCGTG TCTGTCCGGG 3600
TCGTGGCTAT ATCCTTCGTG TCACAGCGTA CAAGGGGGA TATGGAGAAG
GTTGCGTAAC CATTGATGAG ATCGAGAACCA ATACAGACGA ACTGAAGTTT 3700
AGCAACTGCG TAGAAGAGGA AATCTATCCA AATAACACGG TAACGTGTAA
TGATTATACT GTAAATCAAG AAGAATACGG AGGTGCGTAC ACTTCTCGTA 3800
ATCGAGGATA TAACGAAGCT CCTTCCGTAC CAGCTGATTA TGCAGTCAGTC
TATGAAGAAA AATCGTATAC AGATGGACGA AGAGAGAATC CTTGTGAATT 3900
TAACAGAGGG TATAGGGATT ACACGCCACT ACCAGTTGGT TATGTGACAA
AAGAATTAGA ATACTTCCCA GAAACCGATA AGGTATGGAT TGAGATTGGA 4000
GAAACCGGAAG GAACATTAT CGTGGACAGC GTGGAAATTAC TCCTTATGGA
GGAA (end HD-1)

FIG. 11B

M D N N P N I N E C I P Y N C L S N P E V E V L G G E R I E
T G Y T P I D I S L S L T Q F L L S E F V P G A G F V L G L
V D I I W G I F G P S Q W D A F L V Q I E Q L I N Q R I E E
F A R N Q A I S R L E G L S N L Y Q I Y A E S F R E W E A D
P T N P A L R E E M R I Q F N D M N S A L T T A I P L F A V
Q N Y Q V P L L S V Y V Q A A N L H L S V L R D V S V F G Q
R W G F D A A T I N S R Y N D L T R L I G N Y T D Y A V R W
Y N T G L E R V W G P D S R D W V R Y N Q F R R E L T L T V
L D I V A L F P N Y D S R R Y P I R T V S Q L T R E I Y T N
P V L E N F D G S F R G S A Q G I E G S I R S P H L M D I L
N S I T I Y T D A H K G E Y Y W S G H Q I M A S P V G F S G
P E F T F P L Y G T M G N A A P Q Q R I V A Q L G Q G V Y R
T L S S T L Y R R P F N I G I N N Q Q L S V L D G T E F A Y
G T S S N L P S A V Y R K S G T V D S L D E I P P Q N N N V
P P R Q G F S H R L S H V S M F R S G F S N S S V S I I R A
P T F S W Q H R S A E F N N I I P S S Q I T Q I P L T K S T
N L G S G T S V V K G P G F T G G D I L R R T S P G Q I S T
L R V N I T A P L S Q R Y R V R I R Y A S T T N L Q F H T S
I D G R P I N Q G N F S A T M S S G S N L Q S G S F R T V G
F T T P F N F S N G S S V F T L S A H V F N S G N E V Y I D
R I E F V P A E V T F E A E Y D L E R A Q K A V N E L F T S
S N Q I G L K T D V T D Y H I D Q V S N L V E C L S D E F C
L D E K Q E L S E K V K H A K R L S D E R N L L Q D P N F R
G I N R Q L D R G W R G S T D I T I Q G G D D V F K E N Y V
T L L G T F D E C Y P T Y L Y Q K I D E S K L K A Y T R Y Q
L R G Y I E D S Q D L E I Y L I R Y N A K H E T V N V P G T
G S L W P L S A Q S P I G K C G E P N R C A P H L E W N P D
L D C S C R D G E K C A H H S H H F S L D I D V G C T D L N
E D L G V W V I F K I K T Q D G H A R L G N L E F L E E K P
L V G E A L A R V K R A E K K W R D K R E K L E W E T N I V
Y K E A K E S V D A L F V N S Q Y D Q L Q A D T N I A M I H
A A D K R V H S I R E A Y L P E L S V I P G V N A A I F E E
L E G R I F T A F S L Y D A R N V I K N G D F N N G L S C W
N V K G H V D V E E Q N N Q R S V L V L P E W E A E V S Q E
V R V C P G R G Y I L R V T A Y K E G Y G E G C V T I H E I
E N N T D E L K F S N C V E E E I Y P N N T V T C N D Y T V
N Q E E Y G G A Y T S R N R G Y N E A P S V P A D Y A S V Y
E E K S Y T D G R R E N P C E F N R G Y R D Y T P L P V G Y
V T K E L E Y F P E T D K V W I E I G E T E G T F I V D S V
E L L L M E E

FIG. 12